The Intellectual Salmon Run:

Gurney, T.M.

2014

Summary

We understand technologies as the result of knowledge accumulated over time and applied in varied, and sometimes new, forms. Education and practice allow scientists and researchers to understand the phenomena they observe at a fundamental level, and to devise novel methods to apply their understanding of them.

However, the knowledge that is generated in one locale frequently needs to be translated, transferred or transliterated to find meaningful application in another. In other words, in the dynamics of science and technology, the spawning grounds of theory and the hatching grounds of application are divided by an ocean of experience and time - and it is across this ocean we aim to swim. The transfer of knowledge across the metaphorical ocean of experience and time is not radically different from the reality. The end-results of the vast interplay between individuals, firms, universities and environments - be they products, processes or ideas - follow convoluted paths. It takes a concerted effort to follow and trace these paths, be it at the fine-grained level of two individuals communicating, or at the supra-national policy level. There remains uncertainty in the research that has been produced on knowledge transfer in that many questions remain regarding the operationalisation of knowledge transfer. We still do not know what knowledge is transferred, from where and to whom, how exactly the transfer and reception work, and the conditions surrounding the transfer. In addition, this line of questioning is not only of scholarly interest, but also of interest to society in terms of innovation and innovation policy, higher education and science policy. Industry has a vested interest in this, as knowledge transfer between academia and industry provides a significant portion of the inspiration and knowledge they require to produce and develop products and services.

To deduce the processes and mechanisms involved in knowledge transfer, it is necessary to define the three primary aspects of knowledge transfer. The first involves the knowledge itself – how was it generated, how has it developed and how is it primed for transfer. The second involves the ‘sender’ and ‘receiver’ of the information or knowledge – who are they and how has each contributed to the knowledge. And the third involves the environment – how have the conditions surrounding the knowledge facilitated a productive transfer. These aspects form the basis of my primary question wherein: ***What knowledge elements are transferred from academia to industry, how are they transferred, and what factors influence this transfer?***

In researching the precise knowledge elements being transferred, the scientific publications of the person(s) under study must be positively identified as belonging to that individual and not another researcher of the same name. With the rise of the Asian science systems, and the associated low variance in Asian researcher names, this problem is likely to get worse. To tackle this, we strongly require an understanding of the problems related to name ambiguity, plus a reliable and effective process to accurately disambiguate the sometimes vast number of publications.

1 Disambiguation

Automated approaches to disambiguation are necessary and tends to follow either a computer science or a sociological/linguistic approach or a combination of the two. These approaches have

been successful to a degree but most suffer from a common drawback, that of data discarding. For example, studies utilising key words suffer if any records are missing their keywords. Another example is that of using co-author similarity to determine if two records are from the same author. When using co-authors, how do we handle records with only one author i.e. no co-authors? In practice, these records are discarded, to the detriment of the resulting precision and recall of the algorithm. To address the issue of data accuracy, the second sub-question of this thesis is: ***How can we disambiguate researchers with an effective balance between precision and recall?***

Three problems were found with current techniques: discarding of data, limited selection of metadata, and a lack of consideration for evolving research streams and actual contributions to research by authors. The solution to handling the first two problems, data discarding and limited metadata, was to select the best possible alternate combination of available metadata. In other words, for records missing one or more fields, the proportional discriminating powers of the available metadata were adjusted depending on available combinations. All this occurs 'on the fly' within the algorithm, based on previous calibrations. The last problems, that of evolution of terminologies, research streams, or knowledge homogeneity, and author contributions rely on recognising the changing roles, topics and requirements of researchers as they progress through their academic life cycle. For department heads, author status is occasionally 'honorary', in that much of the research and write-up was actually performed by others. In evaluations, however, every author listing counts. In other current techniques, similarity calculations assume input from *all* authors, in *equal* parts. Relying on this assumption leads to networks of publications by authors, exhibiting low or no similarity. This problem is tackled by assuming that contributions by authors vary according to listing (an exception is made for alphabetically ordered listings) and the discriminating powers of the indicators were adjusted to allow for this. For example, the importance of the title or abstract words (most likely selected by the authors writing the paper) is reduced, whereas the journal name (an aspect more likely to be decided on strategic grounds by the head of department) is granted more discriminatory power in the algorithm.

A similar practice is deployed to adjust for the evolution of research streams. Over time, a researcher's usage of words will change to reflect changing fields of interest and publishing. When comparing records published ten years apart, by apportioning less discriminatory power to title or abstract words but more power to journal field characterisation or cited references, one can account for the evolution of researchers' interests. Taking the dynamic selection of alternate metadata together with the adjustments for evolving research interests and contributions means that the developed algorithm allows for more accurate and inclusive results.

2 Knowledge transfer

On a practical level, knowledge transfer and associated mechanisms typically focus on mediums, examples of which include technology or skills where participants receive the knowledge required to perform tasks with a certain technology through the construction and utilisation of that technology itself. Transfer mediums are typically codified in publications and patents or can be tacit. Commonly used indicators of knowledge transfer are based on patent and publication data. Knowledge transfer has typically been addressed in the extant literature as something that occurs as matter of fact. However, there are more complex processes at work within knowledge transfer, other than merely assuming or expecting occurrence. To start, the actual knowledge elements transferred serve as a black box and what is missing is an adequate methodology for *quantifying* the tracks and knowledge being transferred. To aid in answering the primary research question, a second sub questions is necessary, specifically ***how can we identify knowledge elements and their attributes in an operational way, and what elements are transferred between actors?***

To answer this question, a novel method was developed to analyse specific contributions by individuals and institutions, to the developmental routes of specific quanta of knowledge. This method was applied to an individual firm founder, looking at (a) how the scientific background of the patent corpus links to the scientific output of the inventor; (b) how a researcher operates in a collaborative environment, and if the contributions of those contributors are visible in the patent corpus; and (c) whether the inventor demonstrates a level of adaptive knowledge acquisition, necessary for the development of a technology.

The non-patent literature references (NPLR) found in patent applications were used to link patent and publication data. The NPLRs and the author/inventor's corpora of publications were grouped together, linking their shared title words and cited references, and then clustered. The topical foreground and cognitive background of the research cited by the patent applications, and that of the author/inventor's whole corpus of publications, provided a clear view of the required knowledge platforms and absorptive capacity of the inventor and of the contribution it makes to the development and transfer of knowledge

For further clarification, each of the inventor/author's research streams was differentiated further by the introduction of 'concept clusters'. With these concept clusters, the specific, rather than general, contributions made by the researcher involved were identified. This approach took into account the role of co-authors and co-inventors, identifying the specific expertise these collaborators added to the technologies. The institutional affiliation of the inventors and authors also gave an institutional approach to the contributions to the technologies. Additionally, the multidirectional aspect of knowledge and skill transfer between basic and applied research can be examined in minute detail using this method, including the absorptive capacity of individuals and institutions.

The primary results of this show that during the initial stages of a technology's development, the individual in the case study recognised the importance of exogenous knowledge sources. The expertise of his network of co-inventors and co-authors increased his scope and ability to source new knowledge required for the technologies. The technologies were initially based upon the individual's knowledge and skills set, but some aspects were outside his expertise. To combat this perceived gap in knowledge, the assimilation of new knowledge was necessary, and this was ultimately visible in the patent and publication analysis output. Research cited by patent applications that did not fall within the individual's or his co-inventors' expertise was quickly incorporated into the research agenda. The result of this research was seen in the increased publication output in these sub-fields and in some cases the eventual citing of this supplementary research by the patent applications in subsequent developmental stages of the technologies.

3 Absorptive capacity and academic spin-offs

Absorptive capacity may be considered both in terms of the individuals comprising the firm, and as the firm itself. As stated by Cohen and Levinthal, "Beyond diverse knowledge structures, the sort of knowledge that individuals should possess to enhance organizational absorptive capacity is also important. Critical knowledge does not simply include substantive, technical knowledge; it also includes awareness of where useful complementary expertise resides within and outside the organization". In this manner a key aspect is the communication between the firm and the outside world. The concept of absorptive capacity was expanded on to include potential and realised absorptive capacity. These include, for potential absorptive capacity, *acquisition* – which necessitates the taking of stock or inventory of the current assets and knowledge platforms; and *assimilation* - which requires the knowledge intended to be brought in not only to be understood theoretically but also in terms of its place within current knowledge platforms. In realised absorptive capacity, the dimensions of *transformation* – which includes the ability to create novel knowledge by adding external knowledge to the current platform, and *exploitation* – in which results of the combination are brought to light. These could include, but are not limited to, patent applications, scientific publications or new work processes.

To address some of the complex processes in measuring knowledge transfer and absorptive capacity, studies frequently involve academic spin-offs because they provide the clearest identifiable path of knowledge transfer, where an idea can be followed from its inception to its commercial roll-out through a specific individual or group. Spin-offs embody an idea which was developed in academia and deemed to be commercially viable, but they require a dedicated entity to manifest. Overall, studies on spin-offs provide indications of the roles of the individuals involved with the knowledge transfer, as well as the source and end-user environments of the knowledge, but do not examine the effects of the individuals and their environments on the actual knowledge elements being transferred.

Returning to the transfer of knowledge elements, in order to link absorptive capacity and spin-offs, we examine a common route to enabling the infrastructure for absorptive capacity. For spin-offs, the environment is crucial for absorptive capacity to occur. The environment offers firms a *choice* of knowledge, and *access* to an environment is often the first step for firms stepping outside the university. For academic spin-offs, an environment that provides this is often a Science Park.

4 Science Parks

Science Parks provide an environment to promote knowledge transfer and interactions between firms, universities and small labs. They provide a contact space between the 'fast applied science' of industry and the 'slow basic science' of the university and provide a technological platform for economic development at a regional or national level.

Science park locations primarily appeal to firms which are either industry-based spin-outs, or academic spin-offs. There are three distinct reasons at the heart of the motivations of each type of firm to join a Science Park , the first of which is related to neoclassical theory in which transport, labour costs, distance to customers, and agglomeration economies are influential. The second set of reasons stem from behavioural aspects including the presence of mediators, gatekeepers or information channels in the form of the Science Park management. Additionally, the reputational advantages of situating in a Science Park play a large role in influencing firm founders to locate in a park. Most importantly for this thesis, the third set of reasons relate to structuralist approaches, including access to an innovative, networked environment, in which the presence of a Higher Education Institution plays a central role. From this, the third and last sub-question: ***What resources, and from which actors and operational spheres, contribute most significantly to the development of an academic spin-off and its host technology?***

For firms choosing to locate within a Science Park, neoclassical location theory tends to dominate the decision processes of the firm founder. This typically includes logistical issues, such as the proximity to the founder's home. This is not to say that firms interviewed considered only these issues, but rather that *practicalities* won over *potentialities.* In interviews with firm founders, all expressed interest in being able to collaborate with other firms in the Science Park (in line with one of the espoused benefits of locating within a Science Park). However, there was little evidence in the publication and patent data to show that they actually conducted collaborative research with other firms located at the Park.

That is not to say that there was no collaboration, rather there was no substantial evidence of collaboration. From the patent and publication data, the regional and international characteristics of co-assignees and co-authors show that for almost all of the firms, collaborative activities were common, but with firms *outside* the Science Park. Academic collaborations were primarily with the local HEI, Leiden University in this case, and a few founders maintained strong links with their alma maters beyond Leiden. This was reflected in the interview data where the interactions were internal, i.e. initiated before firm's formation and location to the Science Park, and external, i.e. with academic and industrial partners elsewhere in the country and abroad.

Social capital as a resource can be considered supplementary and enabling to the stock knowledge, financial capital and skills of an entrepreneur. For the firm founders who exploited their networks, they drew capital from either internal sources (i.e. historically through personal relationships prior to firm formation) or from sources external to the Science Park. Only a few firms reported any interactions of any nature with the Science Park administration or other firms located at the Science Park. We found that the scientific capabilities of the firm founder were significant in developing and expanding the firm's scientific base, and for its eventual patent output. The substantial similarities between the patent content and the scholarly output of the firm founder (our proxy being the co-location of NPLRs and the founder's publication corpora) and the number of active research streams at and after incorporation both suggest that the scientific base of the founder had a large supporting role.

The lifeblood of a Science Park is its tenants. Science Parks compete for tenants, and those tenants exist in a *coopetitive* environment. Tenants compete for access to networks and the benefits these networks bring. They also are presented with the opportunity to cooperate with other tenants in the park, sharing resources and mitigating risks whilst expanding their potential access to networks. To counteract the diversity issues in a Science Park in terms of evaluating their utility, more emphasis is needed on analysing the knowledge structures and competences of the firms.

5 All together

The overarching research question of this thesis (*what knowledge elements are transferred from academia to industry, how are they transferred, and what factors influence this transfer?)* is at first glance a broad question. It is necessarily broad so as to encompass the complexity of knowledge transfer in relation to absorptive capacity, social capital and the environment in which these knowledge transfer processes take place. In anticipation of the inevitable question of applicability, this thesis goes some way to providing a toolbox for parties that are interested in discovering which elements are transferred, where to and where from, and what factors influence this transfer, for *their* specific field of application.

The underpinnings of this methodological toolbox begin with considering the effect of an individual's previous research on future research plans, and the similarities between past and present current research streams. The methods developed and insights provided in the chapter on disambiguation allow us to analyse factors such as the similarities and differences between research conducted during the PhD phase of a researcher's career and the professorial phase. Over time, an individual's research contributions may change with rank and eventual specialisation, but their incorporated knowledge and skill sets developed remain. Research conducted in academia and eventually applied in industry follows a convoluted path. We needed to gain an understanding of this path for our disambiguation algorithms to succeed, and such an understanding provides a first glimpse at what knowledge elements are transferred over time.

The methodology and case study in the third and fourth chapters serve as a vehicle to examine the specific contributions of an existing knowledge base to the development of a technology platform. In other words, identifying the knowledge elements and, to a certain extent, how they are transferred. The knowledge base does not necessarily come from one individual, but also from co-authors and co-inventors, and from other researchers working in different research settings. The methodology outlined in these chapters provided a toolkit for us to uncover the linkages between research conducted within academia and the eventual application of that research in industry, and the case study provided an example of what our approach can reveal. By applying our new method to a real case study, we demonstrated the ability to combine exogenously generated knowledge with a current knowledge base. New research conducted by Nakamura was guided by previous efforts, indicating that research practices and results are constantly evolving to inform, guide and provide the basis for extensions to different technologies. With detailed descriptions of the linked chains of research, we showed that a research corpus of an individual and their co-inventors and co-authors, can be readily recognised and identified in the exploitation of their research (i.e. in patents). The thematic links between the technologies and the underlying science was clearly identified in this chapter, verifying our methodological approach.

There is a long list of purported benefits for a firm to decide to locate to a Science Park. Its close proximity to a HEI and the prevalence of like-minded firms in the vicinity are examples of the reasons firms decide to locate to a Science Park. In social capital terms, a Science Park provides accessibility to resources, but for the firms in the case study, for the most part, these resources existed *in potentia*. It is important to note that for the firms under study there were no access barriers imposed by the Science Park. All the firm founders considered the Science Park to be an important *potential* source of collaborations and customers. If the opportunity arose, all the firm founders stated they would consider it. However, using the lens of only the purported benefits of locating to a Science Park, we failed to see more than practical benefits.

It is the firms that drive the success of a Science Park, and each firm is driven by its own scientific capacities and potential market linkages. For firms to truly enjoy the network benefits of a Science Park, there should be an overlap of not only their fundamental or applied sciences, but also of potential collaborators and customers. As such, there needs to be a greater emphasis on the underlying sciences and technologies hosted by each firm when setting up a Science Park.

An important step to answering the primary research question was to consider two perspectives: access to resources and technology development. It was necessary to blend the two perspectives, as many of the opportunities in one arise from the other. Ideas generated by a scientist in academia are not initially beholden to entrepreneurial dynamics within their network. They are, however, subject to the incentives structure of academia. New research streams are common in academia, where networks of scientists contribute to one another's research, iteratively guiding the development of an idea or technology. If a certain stream or idea is deemed suitable for exploitation, the scientist/entrepreneur/founder's various networks simultaneously come into play, and the opportunities for further scientific development diminish. In their place, commercialisation of the idea becomes paramount.

Incentive structures shape both the strategies of academic researchers and industrial researchers in terms of valuing their research and results, and thus what aspects of the research will be developed and transferred. For academic spin-offs, access to scientific and technical networks and/or resources is not restricted to academia but extend into industry. The development of these resources is crucial to the development of the technology, but is secondary for the spin-off itself. It is access to the organisational, financial and regulatory networks that tops the priority list and begins to refine the technology. The characteristics of an idea conceived in academia and transferred to industry vary in many ways but, using the developed tools, we can identify its origins, track its evolutionary path, and examine the effects of the environment.